



TITLE:

# A 25/17-Approximation Algorithm for the Stable Marriage Problem with One-Sided Ties

AUTHOR(S):

Iwama, Kazuo; Miyazaki, Shuichi; Yanagisawa,  
Hiroki

---

CITATION:

Iwama, Kazuo ...[et al]. A 25/17-Approximation Algorithm for the Stable Marriage Problem with One-Sided Ties. *Algorithmica* 2014, 68(3): 758-775

ISSUE DATE:

2014-03

URL:

<http://hdl.handle.net/2433/226941>

RIGHT:

The final publication is available at Springer via <https://doi.org/10.1007/s00453-012-9699-2>; The full-text file will be made open to the public on 01 March 2015 in accordance with publisher's 'Terms and Conditions for Self-Archiving'; この論文は出版社版ではありません。引用の際には出版社版をご確認ご利用ください。; This is not the published version. Please cite only the published version.

<b>Algorithmica manuscript No.</b> (will be inserted by the editor)
--

---

## A 25/17-Approximation Algorithm for the Stable Marriage Problem with One-Sided Ties

Kazuo Iwama · Shuichi Miyazaki ·  
Hiroki Yanagisawa

Received: date / Accepted: date

**Abstract** The problem of finding a largest stable matching where preference lists may include ties and unacceptable partners (MAX SMTI) is known to be NP-hard. It cannot be approximated within  $33/29$  ( $> 1.1379$ ) unless  $P=NP$ , and the current best approximation algorithm achieves the ratio of 1.5. MAX SMTI remains NP-hard even when preference lists of one side do not contain ties, and it cannot be approximated within  $21/19$  ( $> 1.1052$ ) unless  $P=NP$ . However, even under this restriction, the best known approximation ratio is still 1.5. In this paper, we improve it to  $25/17$  ( $< 1.4706$ ).

**Keywords** the stable marriage problem · the stable marriage with ties and incomplete lists · approximation algorithm · integer program · linear program relaxation · integrality gap

---

This work was supported by JSPS KAKENHI Grant Number 22240001, 20700009, 24500013.

---

Kazuo Iwama  
Graduate School of Informatics, Kyoto University,  
Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan.  
Tel.: +81-75-753-5372  
Fax: +81-75-753-5972  
E-mail: iwama@kuis.kyoto-u.ac.jp

Shuichi Miyazaki  
Academic Center for Computing and Media Studies, Kyoto University,  
Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan.  
Tel.: +81-75-753-7418  
Fax: +81-75-753-7440  
E-mail: shuichi@media.kyoto-u.ac.jp

Hiroki Yanagisawa  
IBM Research - Tokyo,  
NBF Toyosu Canal Front Building, 6-52, Toyosu 5-chome, Koto-ku, Tokyo 135-8511, Japan.  
Tel.: +81-3-5144-2932  
Fax: +81-3-3532-0061  
E-mail: yanagis@jp.ibm.com

## 1 Introduction

The *stable marriage problem* [3, 5] is a classical bipartite matching problem. In its original setting, an instance consists of  $n$  men,  $n$  women, and each person's preference list, where a preference list is a totally ordered list of all the members of the opposite sex according to his/her preference. A *matching* is a set of disjoint pairs each consisting of a man and a woman. For a matching  $M$ , a pair of a man  $m$  and a woman  $w$  is called a *blocking pair* if both prefer each other to their current partners. A matching with no blocking pair is called *stable*. Gale and Shapley [3] showed that every instance admits at least one stable matching, and proposed an  $O(n^2)$ -time algorithm, known as the *Gale-Shapley algorithm*, to find one.

There are several examples of using the stable marriage problem in assignment systems, including residents/hospitals matching [5, 19], students/schools matching [1, 2, 22], etc. Clearly, the restrictions that each preference list must be *strict* (i.e., it is a totally ordered list) and *complete* (i.e., it includes all the members of the opposite side) are unrealistic for such applications. Two natural relaxations are then to allow for *ties* and *incompleteness*. (There are three definitions of stability when ties are allowed [5, 8]. In this paper, we consider *weak stability*, which is the most natural notion among the three.) Applying either one or both of these extensions does not affect the validity of the properties that a stable matching exists for any instance and that one can be found in polynomial time. Hence a stable matching can be found efficiently even with these extensions.

However, if one is concerned with the *size* of stable matchings, the situation changes significantly. In the original setting and an extension with only ties, a stable matching is a perfect matching by definition. In an extension with only incomplete preference lists, a stable matching may no longer be perfect but for a fixed instance, all of the stable matchings have the same size due to the famous Rural Hospitals Theorem [4]. Thus the problem of finding a largest stable matching is solvable in polynomial time for all of these cases. In contrast, if we allow *both* extensions, one instance can have stable matchings of different sizes, and hence the problem of finding a largest stable matching (which we call hereafter *MAX SMTI* (MAXimum Stable Marriage with Ties and Incomplete lists)) is no longer trivial. In fact, this problem was shown to be NP-hard [11, 17]. Since a stable matching is a maximal matching, any two stable matchings differ in size by at most a factor of two. Hence, constructing a polynomial-time 2-approximation algorithm is easy.

There has been a sequence of improvements of the 2-approximation ratio. The first attempt was made by using a local search type algorithm, which successively increases the size of a stable matching at each iteration [12, 13], but these upper bounds approach 2 as  $n$  goes to infinity. The first upper bound strictly better than 2 was obtained along this line. Iwama et al. [14] obtained an upper bound of 1.875 by modifying the aforementioned local search. Later, Király [15] improved it to  $5/3$  by using a completely different idea. His algorithm is a modification of the Gale-Shapley algorithm so that each

man may propose to the same woman more than once, and the roles of men and women are exchanged during the execution. McDermid [18] improved Király's algorithm by exploiting a classical result in graph theory, the Gallai-Edmonds decomposition, and obtained an upper bound of 1.5, which is the current best approximation ratio. Meanwhile, the best-known lower bound is  $33/29$  ( $> 1.1379$ ) under the assumption that  $P \neq NP$  [23].

For a given instance, let  $M_{opt}$  be a largest stable matching and  $M$  be any stable matching. Consider a union of  $M_{opt}$  and  $M$ , which can be seen as a bipartite graph. Each connected component is an alternating path or cycle. If every connected component is a path of length three that contains two  $M_{opt}$ -edges and one  $M$ -edge, then  $|M_{opt}| = 2|M|$ , which is the worst possible case mentioned above. All of the above approximation algorithms were designed to exclude as many such length-three paths as possible. The ratio of 1.875 [14] was achieved by excluding a constant fraction (relative to the size of  $M$ ) of such components. And in [15], beyond his  $5/3$  result, Király was also able to remove all such length-three paths (which led to the approximation ratio of 1.5) if the instance has ties *on only one side*, which is another main result of that paper. As an extension, McDermid [18] finally succeeded in excluding them completely for general instances.

A natural extension of this line of research is to attack augmenting paths of *length five*. In order to break the bound of 1.5, we need to remove those paths at least by a constant fraction. This is apparently a challenging goal for general instances. Even if ties are on only one side, Király stated in Conjecture 3 in [15] that breaking the 1.5-approximation implies something “surprising,” that is, breaking the 1.5-approximation upper bound of MAX SMTI with one-sided ties implies breaking the 3-approximation upper bound of the minimum vertex cover problem in 3-uniform hyper-graphs (which disproves the Unique Games Conjecture (UGC) [16]).

**Our Contributions.** In this paper, we improve the approximation ratio from 1.5 to  $25/17$  ( $< 1.4706$ ) for instances with one-sided ties, which disproves Király's conjecture under UGC. We note that this approximation ratio also holds for the hospitals/residents problem (i.e., many-one variant) with one-sided ties (see [9] for the relationship between approximability of the stable marriage problem and the hospitals/residents problem).

The basic idea is to use an integer program (IP) and its linear program (LP) relaxation, which is summarized as follows: note that the Gale-Shapley algorithm (GS) consists of a sequence of proposals from men to women (see [5] for details). Király's algorithm, GSA1, is similar to GS, but is different in that each man goes through his list twice for proposals. In GSA1, each man has one of two possible states, “unpromoted” or “promoted” (these terms are taken from McDermid's paper [18]). Each man is initially unpromoted, and once he has proposed to all of the women on his list, he becomes promoted and starts again by making proposals from the top of the list. When a woman receives proposals from two men with the same preference but different states, she always selects the promoted one. It should be more powerful to use not

only two different states, unpromoted and promoted, but more quantitative information for the same purpose. Our new idea is to formulate MAX SMTI as an IP by generalizing the IP formulation for the original stable marriage problem [20,21], and then to solve its LP relaxation (in polynomial time). We use this optimal solution to define the state of each man.

Finally, we briefly discuss the limits of our approach using the integrality gap of our IP formulation.

**Related Results.** As mentioned above, general MAX SMTI is approximable within 1.5, but cannot be approximated with a ratio smaller than  $33/29$  ( $> 1.1379$ ) unless  $P=NP$  and smaller than  $4/3$  ( $> 1.3333$ ) under UGC [23]. If the length of each man's list is bounded by 2, it is solvable in polynomial time even if women's lists are of arbitrary length, while it is NP-hard even if the length of each preference list is bounded by 3 [10]. The only known approximability result with a ratio smaller than 1.5 is a randomized approximation algorithm that achieves an expected approximation ratio of  $10/7$  ( $< 1.4286$ ) for the special case that ties appear on one side only and the length of each tie is at most two [6].

The restriction for MAX SMTI that ties appear in preference lists of one sex only is quite natural in practice. For example, it is reported that in the Scottish Foundation Allocation Scheme (SFAS), a hospitals/residents matching system in Scotland, residents are required to submit strictly ordered preference lists while each hospital's list may contain one tie [9]. Even under this restriction, MAX SMTI remains NP-hard and is not approximable with a ratio smaller than  $21/19$  ( $> 1.1052$ ) unless  $P=NP$ , and smaller than  $5/4$  under UGC [7]. In this restricted setting, Irving and Manlove [9] presented a  $5/3$ -approximation algorithm. Király [15] improved it to 1.5, which is the previous best upper bound as well as the McDermid bound [18] for the general case.

## 2 Preliminaries

An instance  $I$  of MAX SMTI comprises  $n$  men,  $n$  women and each person's preference list that may be incomplete and may include ties. If a person  $p$  includes a person  $q$  (of the opposite sex) in  $p$ 's preference list, we say that  $q$  is *acceptable* to  $p$ . Without loss of generality, we assume that  $m$  is acceptable to  $w$  if and only if  $w$  is acceptable to  $m$ . A matching  $M$  is a set of pairs  $(m, w)$  such that  $m$  is acceptable to  $w$  and vice versa, and each person appears at most once in  $M$ . If  $(m, w) \in M$ , we say that  $m$  ( $w$ ) is *matched* in  $M$ , and write  $M(m) = w$  and  $M(w) = m$ . If a person  $p$  does not appear in  $M$ , we say that  $p$  is *single* in  $M$ .

If  $m$  strictly prefers  $w_i$  to  $w_j$ , we write  $w_i \succ_m w_j$ . If  $w_i$  and  $w_j$  are tied in  $m$ 's list (including the case that  $w_i = w_j$ ), we write  $w_i =_m w_j$ . The statement  $w_i \succeq_m w_j$  is true if and only if  $w_i \succ_m w_j$  or  $w_i =_m w_j$ . We use similar notations for women's preference lists. We say that  $m$  and  $w$  form a *blocking pair* for a matching  $M$  (or simply,  $(m, w)$  *blocks*  $M$ ) if the following three conditions are met: (i)  $M(m) \neq w$  but  $m$  and  $w$  are acceptable to each other,

(ii)  $w \succ_m M(m)$  or  $m$  is single in  $M$ , and (iii)  $m \succ_w M(w)$  or  $w$  is single in  $M$ . A matching  $M$  is called *stable* if there is no blocking pair for  $M$ . MAX SMTI is the problem of finding a largest stable matching.

The *approximation ratio* of an approximation algorithm  $T$  is  $\max\{opt(I)/T(I)\}$  over all instances  $I$ , where  $opt(I)$  and  $T(I)$  are the sizes of the optimal and the algorithm's solutions, respectively.

The following IP formulation of MAX SMTI instance  $I$ , denoted by  $IP(I)$ , is a generalization of the one for the original stable marriage problem given in [20, 21]. For each (man, woman) pair  $(m, w)$ , we introduce a variable  $x_{m,w}$ .

$$\text{Maximize:} \quad \sum_i \sum_j x_{i,j}$$

$$\text{Subject to:} \quad \sum_i x_{i,w} \leq 1 \quad \forall w \quad (1)$$

$$\sum_j x_{m,j} \leq 1 \quad \forall m \quad (2)$$

$$\sum_{j \succeq_m w} x_{m,j} + \sum_{i \succeq_w m} x_{i,w} - x_{m,w} \geq 1 \quad \forall (m, w) \in A \quad (3)$$

$$x_{m,w} = 0 \quad \forall (m, w) \notin A \quad (4)$$

$$x_{m,w} \in \{0, 1\} \quad \forall (m, w) \quad (5)$$

Here,  $A$  is the set of mutually acceptable pairs, that is,  $(m, w) \in A$  if and only if each of  $m$  and  $w$  includes the other in the preference list. In this formulation, “ $x_{m,w} = 1$ ” is interpreted as “ $m$  and  $w$  are matched,” and “ $x_{m,w} = 0$ ” otherwise. Thus the objective function is equal to the size of a matching. Note that Constraint (3) ensures that  $(m, w)$  is not a blocking pair. When  $x_{m,w} = 1$ , all three terms of the lefthand side are 1 and hence Constraint (3) is satisfied. When  $x_{m,w} = 0$ , either the first or the second term of the lefthand side must be 1, which implies that  $m$  (respectively  $w$ ) must be matched with a partner as good as  $w$  (respectively  $m$ ).  $LP(I)$  denotes the linear program relaxation of  $IP(I)$  in which Constraint (5) is replaced by “ $0 \leq x_{m,w} \leq 1$ .”

### 3 Approximation Algorithm

In the following, we assume that the men's lists are strict and the women's lists may contain ties. Algorithm 1 gives the pseudo-code for our algorithm GSA-LP. We use a variable  $f(m)$ , which assigns a non-negative value to each man  $m$ . The value of  $f(m)$  is initially set to zero but increases as the algorithm proceeds. We also use another variable  $p(m)$ , which stores the current position for  $m$  in his preference list. When man  $m$  proposes to woman  $w$ ,  $w$  accepts this proposal if either (a)  $w$  is single in  $M$ , (b)  $m \succ_w M(w)$ , or (c)  $m =_w M(w)$  and  $f(m) > f(M(w))$ . Otherwise,  $w$  rejects  $m$ 's proposal. When  $w$  accepts  $m$ 's proposal, we let  $M := M \cup \{(m, w)\}$  for Case (a), and let  $M := M \cup \{(m, w)\} \setminus \{(M(w), w)\}$  for Cases (b) and (c).

---

**Algorithm 1** GSA-LP (Gale-Shapley Algorithm with LP solution)

---

**Input:** An SMTI instance  $I$

**Output:** A matching  $M$

```

1: Formulate the given instance  $I$  as an integer program  $IP(I)$ 
2: Solve its LP relaxation  $LP(I)$  and obtain an optimal solution  $x^*(= \{x_{i,j}^*\})$ 
3: Let  $M := \emptyset$ 
4: Set  $f(m) := 0$  and  $p(m) := 1$  for each man  $m$ 
5: while there exists a man  $m$  such that ( $m$  is single in  $M$ ) and ( $f(m) \leq 3$ ) do
6:   Let  $m$  be an arbitrary such man
7:   if  $p(m)$  is larger than the length of  $m$ 's preference list then
8:     if  $f(m) \leq 1$  then
9:       Set  $f(m) := 2$  and  $p(m) := 1$ 
10:    else
11:      Set  $f(m) := f(m) + 1$  and  $p(m) := 1$ 
12:    end if
13:  else
14:    Let  $w$  be the  $p(m)$ -th woman of  $m$ 's preference list
15:    if  $m$  has not proposed to  $w$  yet then
16:      Set  $f(m) := f(m) + x_{m,w}^*$ 
17:      Let  $m$  propose to  $w$ 
18:      Set  $p(m) := 1$ 
19:    else
20:      Let  $m$  propose to  $w$ 
21:      Set  $p(m) := p(m) + 1$ 
22:    end if
23:  end if
24: end while
25: return  $M$ 

```

---

Here is an intuition of our new algorithm GSA-LP: It basically consists of three rounds of proposal sequences by each man  $m$  with the rank-adjusting value  $f(m)$ . In the first round,  $f(m)$  is increased by  $x_{m,w}^*$  whenever  $m$  sends a proposal to  $w$  for the first time (Lines 15–18). The key thing here is that if  $m$  is rejected by this new woman  $w$  (either immediately or later after once accepted), then he restarts his sequence of proposals from the top of his list (Line 18). Note that in the restarted sequence of proposals, women he proposes to are not new until  $w$ . Up to that woman,  $f(m)$  does not change and the restart does not happen.

If  $m$  has proposed to all of the women in his list and he is still single, then the value of  $f(m)$  is set to 2 (Lines 8–9) and  $m$  goes to the second round. In the second round, the value of  $f(m)$  does not change and  $m$  sends a sequence of proposals from top of his list again. If  $m$  is still single after finishing this sequence, then  $f(m)$  is increased by 1 (namely,  $f(m)$  is set to 3) (Lines 10–11) and  $m$  restarts another sequence of proposals again (the third round). If  $m$  is still single after the third round,  $f(m)$  is set to 4 and hereafter  $m$  sends no more proposals. Note that not every man goes into the third round, but if a man enters the third round, then it is guaranteed that his second round is completed.

Note that just before  $m$  first proposes to  $w$  in his first round, he has proposed to all of the preceding women with the value  $f(m) = \sum_{j \succ_m w} x_{m,j}^*$

but was rejected by all of them. Also, when a man  $m$  proposes to the woman at the tail of his list for the first time,  $f(m) \leq 1$ . Finally, note that if  $m$  is single in  $M$  at the termination of GSA-LP, he has proposed to all the women in his list with the value  $f(m) = 3$ , and  $f(m) = 4$  holds at the termination of GSA-LP.

It is not hard to see that GSA-LP runs in polynomial time. For the correctness, let  $M$  be a matching obtained by GSA-LP, and let  $(m, w) \in A \setminus M$ . If  $m$  has not proposed to  $w$ , then  $M(m) \succ_m w$  and hence  $(m, w)$  is not a blocking pair. If  $m$  has proposed to  $w$ , then  $m$  was rejected by  $w$  at some time, and at this moment  $w$  must have been matched with a man  $m'$  such that  $m' \succeq_w m$ . Since hereafter  $w$  never accepts a proposal from a man inferior to  $m'$ , it must be the case that  $M(w) \succeq_w m' \succeq_w m$ . Hence  $(m, w)$  is not a blocking pair, so that  $M$  is stable.

The following instance  $I_1$ , presented in [15], is a worst-case instance for both the Király [15] and the McDermid [18] algorithms.

$$\begin{array}{ll} m_1: & w_1 \\ m_2: & w_2 \ w_1 \\ m_3: & w_2 \ w_3 \end{array} \qquad \begin{array}{ll} w_1: & m_2 \ m_1 \\ w_2: & (m_2 \ m_3) \\ w_3: & m_3 \end{array}$$

Here, when two or more men are in the same parenthesis in a woman  $w$ 's preference list, it means that they are tied in  $w$ 's list. In the current example,  $m_2$  and  $m_3$  are tied in  $w_2$ 's list. The optimal solution of  $I_1$  is  $M_{opt} = \{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ , but these two algorithms may output a stable matching  $M = \{(m_2, w_1), (m_3, w_2)\}$ . During the execution of these algorithms,  $w_2$  receives proposals from both  $m_2$  and  $m_3$ . To obtain  $M_{opt}$ , we want to make  $w_2$  choose  $m_2$  rather than  $m_3$ , so that  $m_3$  proposes to  $w_3$ . But in these algorithms, there are executions in which  $w_2$  chooses  $m_3$  (depending on the order of the proposals).

In contrast, our algorithm GSA-LP always obtains the optimal solution. The LP relaxation  $LP(I_1)$  has the following (unique) optimal solution  $x^*$ .

- $x_{m_i, w_j}^* = 1$  for  $i = j$ ,
- $x_{m_i, w_j}^* = 0$  for  $i \neq j$ .

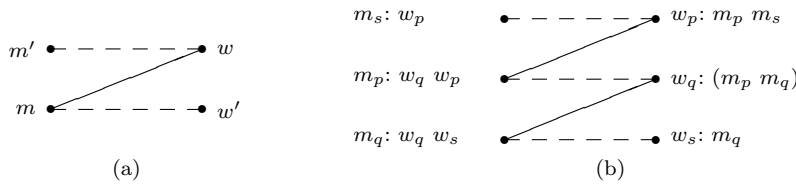
Hence in GSA-LP, when  $m_2$  proposes to  $w_2$  for the first time,  $f(m_2)$  is already 1, but  $f(m_3) = 0$  unless  $m_3$  proposes to  $w_3$ . Therefore,  $w_2$  always chooses  $m_2$  and hence we can obtain  $M_{opt}$ .

## 4 Analysis of the Approximation Ratio

### 4.1 Overview of the Analysis

Let us fix an instance  $I$ . Let  $M_{opt}$  be an optimal solution, namely one of the maximum stable matchings of  $I$ , and  $M$  be the stable matching output by GSA-LP. Let  $G = (U, V, E)$  be the bipartite graph obtained by superimposing  $M_{opt}$  on  $M$ , that is, each vertex in  $U$  corresponds to each man, and each vertex





**Fig. 1** Illustrations of (a) an augmenting three-path and (b) an augmenting five-path. Solid lines represent  $M$ -edges and dashed lines represent  $M_{opt}$ -edges.

in  $V$  corresponds to each woman. For simplicity, we do not distinguish between a person and a corresponding vertex (e.g., a vertex of  $G$  corresponding to a man  $m$  is also denoted  $m$ ). If  $m \in U$  and  $w \in V$  are matched in  $M$ , then  $E$  includes an edge  $(m, w)$ , called an  $M$ -edge. Similarly, if  $m$  and  $w$  are matched in  $M_{opt}$ , then  $E$  includes an  $M_{opt}$ -edge  $(m, w)$ . Graph  $G$  contains parallel edges  $(m, w)$  if  $(m, w)$  is a pair in both  $M$  and  $M_{opt}$ . Note that each vertex in  $G$  has degree at most two, and hence each connected component of  $G$  is either a single vertex, an alternating path, or an alternating cycle. In this paper, we refer to a path of length three starting from and ending with  $M_{opt}$ -edges as an *augmenting three-path* (see Fig. 1(a)). We can prove Proposition 1 by following a similar analysis to Király's GSA1 [15].

**Proposition 1** *There is no augmenting three-path in  $G$ .*

*Proof* Suppose that there is such a path  $m'-w-m-w'$ , where  $(m, w)$  is a pair in  $M$  and  $(m', w)$  and  $(m, w')$  are pairs in  $M_{opt}$ . (See Fig. 1(a).) Since the men's lists do not include ties,  $w$  is better than  $w'$  in  $m$ 's list (otherwise,  $(m, w')$  is a blocking pair in  $M$ ). This implies  $m' =_w m$  since otherwise, there exists a blocking pair either in  $M$  or in  $M_{opt}$ . Since  $w'$  is single in  $M$ ,  $m$  never proposed to  $w'$ , meaning  $f(m) \leq 1$  at the end (and hence at any moment) of the algorithm. Since  $m'$  is single in  $M$ ,  $m'$  must send a proposal to  $w$  with  $f(m') = 3$ . But this is a contradiction since  $w$  selected  $m$  rather than  $m'$ .  $\square$

Let  $S$  be the set of men and women who are single in  $M$ . Also, let us partition  $M$  into  $P$ ,  $Q$ , and  $R$  using the graph  $G$ : Consider a path of length five starting from and ending with  $M_{opt}$ -edges, which we will refer to as an *augmenting five-path* hereafter. (See Fig. 1(b). The preference structure shown here will be established later by Lemma 5.) Let  $m_s, w_p, m_p, w_q, m_q$ , and  $w_s$  be the men and the women on this path appearing in this order, that is, both  $m_s$  and  $w_s$  are in  $S$ , and  $w_p = M_{opt}(m_s)$ ,  $m_p = M(w_p)$ ,  $w_q = M_{opt}(m_p)$ ,  $m_q = M(w_q)$ , and  $w_s = M_{opt}(m_q)$ . Let  $P$  be the set of pairs  $(m_p, w_p)$  and  $Q$  be the set of pairs  $(m_q, w_q)$  on all the augmenting five-paths. Let  $R = M \setminus (P \cup Q)$ .

Note that  $G$  does not contain a path of length one (i.e., an isolated edge) as a connected component, because if such a path (e.g., an  $M_{opt}$ -edge  $(m, w)$ ) exists, then  $(m, w)$  is a blocking pair for  $M$ . Graph  $G$  does not contain any augmenting three-path by Proposition 1. An augmenting five-path contains one edge from each of  $P$  and  $Q$ , and three  $M_{opt}$ -edges. Hence the total number of  $M_{opt}$ -edges contained in augmenting five-paths is exactly  $\frac{3}{2}(|P| + |Q|)$ . For

other connected components of  $G$ , the ratio of the number of  $M_{opt}$ -edges to the number of  $M$ -edges is at most  $\frac{4}{3}$  (for augmenting seven-paths). We thus have Lemma 1, which is one way of bounding  $|M_{opt}|$  by  $|P|$ ,  $|Q|$ , and  $|R|$ .

**Lemma 1**  $|M_{opt}| \leq \frac{3}{2}(|P| + |Q|) + \frac{4}{3}|R|$ .

Since  $|M| = |P| + |Q| + |R|$ , Lemma 1 gives  $|M_{opt}| \leq \frac{3}{2}|M| - \frac{1}{6}|R|$ . However this guarantees only the ratio of 1.5 when  $|R| = 0$ , which is exactly the worst-case example for Király's GSA1 [15].

The advantage of our new algorithm is that it allows us to apply another formula to bound  $|M_{opt}|$ , in the following way: Recall that  $x_{i,j}^*$  is the value of  $x_{i,j}$  for the optimal solution  $x^*$  of  $LP(I)$ . Note that if  $x_{m,w}^* > 0$  for  $m, w \in S$ , then  $(m, w) \in A$  by Constraint (4) of  $LP(I)$ , so  $(m, w)$  is a blocking pair for  $M$ , a contradiction. Hence  $\sum_{i,j \in S} x_{i,j}^* = 0$ . Now, let us define the value  $x^*(X)$  for a subset  $X \subseteq M$  as:

$$x^*(X) = \sum_{(m,w) \in X} \left( \sum_j x_{m,j}^* + \sum_i x_{i,w}^* + \sum_{j \in S} x_{m,j}^* + \sum_{i \in S} x_{i,w}^* \right).$$

It is not hard to see that  $x^*(P) + x^*(Q) + x^*(R) = 2 \sum_i \sum_j x_{i,j}^*$ , since we already know  $\sum_{i,j \in S} x_{i,j}^* = 0$ . Hence the optimal value of the objective function of  $LP(I)$  can be written as  $(x^*(P) + x^*(Q) + x^*(R))/2$  and we have that  $|M_{opt}| \leq (x^*(P) + x^*(Q) + x^*(R))/2$ . We will later prove the following key lemma.

**Lemma 2**  $x^*(P) + x^*(Q) + x^*(R) \leq \frac{2}{5}(7|P| + 7|Q| + 9|R|)$ .

Hence we have that  $|M_{opt}| \leq \frac{1}{5}(7|P| + 7|Q| + 9|R|)$ , that is,  $5|M_{opt}| \leq 7(|P| + |Q|) + 9|R|$ . By Lemma 1, we have  $12|M_{opt}| \leq 18(|P| + |Q|) + 16|R|$ . By adding these two inequalities, we have  $17|M_{opt}| \leq 25(|P| + |Q| + |R|) = 25|M|$ . Thus we obtain the following theorem:

**Theorem 1** *The approximation ratio of GSA-LP is at most 25/17.*

## 4.2 Proof of Lemma 2

To simplify the notation, let us define  $\alpha$  and  $\beta$  as:

$$\alpha = \sum_{(m,w) \in P} \sum_{M_{opt}(m) \succeq_m j, j \notin S} x_{m,j}^* \text{ and } \beta = \sum_{(m,w) \in Q} \sum_{M_{opt}(m) \succeq_m j, j \notin S} x_{m,j}^*.$$

Then we can show the following two lemmas, whose proofs will be given later.

**Lemma 3**  $x^*(P) + x^*(Q) + x^*(R) \leq 2|P| + 2|Q| + 6|R| + 4\alpha + 4\beta$ .

**Lemma 4**  $x^*(P) + x^*(Q) + x^*(R) \leq 3|P| + 3|Q| + 3|R| - \alpha - \beta$ .

Using these two lemmas, we prove Lemma 2.

*Proof* By multiplying the inequality of Lemma 4 by four and then adding to the inequality of Lemma 3, we obtain this lemma.  $\square$

The rest of this section is devoted to the proofs of Lemmas 3 and 4. In Sec. 4.3, we give some properties of  $P$ ,  $Q$ ,  $R$ , and the optimal LP solution. In Sec. 4.4, we give upper bounds on  $x^*(P)$ ,  $x^*(Q)$ , and  $x^*(R)$ . Finally we prove Lemma 3 in Sec. 4.5 and Lemma 4 in Sec. 4.6.

#### 4.3 Properties of $P$ , $Q$ , $R$ , and LP Solution

We first show several basic properties that will be used later. Lemma 6 is especially important since it makes a difference between our new algorithm and Király's.

**Lemma 5** *Consider an augmenting five-path in graph  $G$ . Let  $m_s, w_p, m_p, w_q, m_q$ , and  $w_s$  be the men and women on this path appearing in this order (note that  $(m_p, w_p) \in P$  and  $(m_q, w_q) \in Q$ ). Then the following (i) – (v) hold: (i)  $w_q \succ_{m_q} w_s$ , (ii)  $m_p \succ_{w_p} m_s$ , (iii)  $w_q \succ_{m_p} w_p$ , (iv)  $m_p =_{w_q} m_q$ , and (v)  $f(m_p) \leq 1$  and  $f(m_q) \leq 1$  at the termination of GSA-LP.*

*Proof* If  $w_s \succ_{m_q} w_q$ , then  $(m_q, w_s)$  blocks  $M$ . Hence (i) holds (note that the men's lists do not contain ties). Since  $w_s$  is single in  $M$ ,  $m_q$  did not propose to  $w_s$ , and hence  $f(m_q) \leq 1$  (the latter of (v)).

For the same reason as for (i), we have that  $m_p \succeq_{w_p} m_s$ . To show that (ii) holds, we show that  $m_p =_{w_p} m_s$  does not hold. Suppose that this condition holds. Since  $m_s \in S$ ,  $m_s$  must have proposed to  $w_p$  with the value  $f(m_s) = 3$  in the final round. Then, since  $w_p$  is matched with  $m_p$  in  $M$ ,  $f(m_p) \geq 3$  at the termination. Hence  $m_p$  must have completed the second round and hence must have proposed to  $w_q$  with the value  $f(m_p) \geq 2$ . As we have proved above,  $f(m_q) \leq 1$ , but  $w_q$  is eventually matched in  $M$  with  $m_q$ , so  $m_q \succ_{w_q} m_p$ . But then by (i),  $(m_q, w_q)$  blocks  $M_{opt}$ , a contradiction. Hence (ii) holds.

Condition (iii) holds since if not, then  $(m_p, w_p)$  blocks  $M_{opt}$ , and condition (iv) holds since if not, then either  $(m_q, w_q)$  blocks  $M_{opt}$  or  $(m_p, w_q)$  blocks  $M$ .

Suppose that  $f(m_p) > 1$ . Then  $m_p$  must have proposed to  $w_p$  in the second round and hence also to  $w_q$  (recall that  $w_q \succ_{m_p} w_p$ ) with this  $f$  value. But this contradicts the fact that  $m_p =_{w_q} m_q$  and  $w_q$  selected  $m_q$  whose  $f$  value  $\leq 1$ , which completes the proof of (v).  $\square$

Next, we prove several equations: The first ones are

$$\sum_{j \succ_m w, j \in S} x_{m,j}^* = 0 \text{ and } \sum_{i \succ_w m, i \in S} x_{i,w}^* = 0 \text{ for } (m, w) \in M. \quad (6)$$

For the left equation, suppose that there is a woman  $j$  such that  $j \in S$  and  $j \succ_m w$ . Then woman  $j$  must include  $m$  in her list. Hence  $(m, j)$  blocks  $M$ ,

which contradicts the stability of  $M$ . The right equation can also be validated in a similar way. The next equation is

$$\sum_{i=_w m, i \in S} x_{i,w}^* = 0 \quad \text{for } (m, w) \in P \cup Q. \quad (7)$$

Suppose that there exists a man  $i$  such that  $i =_w m$  and  $i \in S$ . Then man  $i$  must include  $w$  in his list. Since  $i \in S$ ,  $i$  should have proposed to  $w$  with  $f(i) = 3$ . This contradicts the fact that  $w$  selected  $m$  whose  $f$  value is at most 1 by Lemma 5(v). Thus this equation holds.

#### Lemma 6

$$\sum_{(m,w) \in P} \sum_{j \succ_m w, j \notin S} x_{m,j}^* \leq \sum_{(m,w) \in Q} \sum_{j \succ_m M_{opt}(m), j \notin S} x_{m,j}^*.$$

*Proof* Consider an augmenting five-path of Fig. 1(b). Note that both  $m_p$  and  $m_q$  obtained their partners in the first round (by Lemma 5(v)). Let  $f(m_q)$  be the final  $f$ -value of  $m_q$ , and  $f'(m_p)$  and  $f'(m_q)$  be the maximum  $f$ -values of  $m_p$  and  $m_q$  respectively, when they proposed to  $w_q$  (note that they may propose to  $w_q$  several times). Then it turns out that

$$\begin{aligned} \sum_{j \succ_{m_p} w_p, j \notin S} x_{m_p,j}^* &= \sum_{j \succ_{m_p} w_p} x_{m_p,j}^* \leq f'(m_p) \leq f'(m_q) \\ &\leq f(m_q) \leq \sum_{j \succ_{m_q} M_{opt}(m_q), j \notin S} x_{m_q,j}^* \end{aligned}$$

for the reasons given shortly. Note that the lemma is proved if we add this inequality for all augmenting five-paths.

Now, here are reasons for the validity of the above inequalities. (i) The first equality is due to Equation (6). (ii) The first inequality is the key fact obtained from our new algorithm: When  $m_p$  proposes to  $w_p$  for the first time, the  $f$ -value must be at least  $\sum_{j \succ_{m_p} w_p} x_{m_p,j}^*$ . Since  $w_q \succ_{m_p} w_p$ ,  $m_p$  must have already proposed to  $w_q$  with at least this value. (iii) The second inequality is due to the fact that  $w_q$  selected  $m_q$  rather than  $m_p$  as her partner (note that  $m_p =_{w_q} m_q$  by Lemma 5(iv)). (iv) The third inequality is obvious since the value of  $f(m_p)$  is monotone non-decreasing. (v) The final inequality is because  $m_q$  proposed neither to  $M_{opt}(m_q)$  (since  $M(m_q) = w_q \succ_{m_q} M_{opt}(m_q)$  by Lemma 5(i)), nor to any woman in  $S$  (because once a woman receives a proposal, she is matched at the end).  $\square$

#### 4.4 Bounding $x^*(P)$ , $x^*(Q)$ , and $x^*(R)$

Let us define the following values.

$$p_m = \sum_{(m,w) \in P} \sum_{j \notin S} x_{m,j}^*, \quad p_w = \sum_{(m,w) \in P} \sum_{i \notin S} x_{i,w}^*, \quad q_m = \sum_{(m,w) \in Q} \sum_{j \notin S} x_{m,j}^*,$$

$$q_w = \sum_{(m,w) \in Q} \sum_{i \notin S} x_{i,w}^*, \quad r_m = \sum_{(m,w) \in R} \sum_{j \notin S} x_{m,j}^*, \quad \text{and} \quad r_w = \sum_{(m,w) \in R} \sum_{i \notin S} x_{i,w}^*.$$

By the definition of  $x^*(X)$  and Constraints (1) and (2), we have

$$\begin{aligned} x^*(P) &= \sum_{(m,w) \in P} \left( \sum_j x_{m,j}^* + \sum_i x_{i,w}^* + \sum_{j \in S} x_{m,j}^* + \sum_{i \in S} x_{i,w}^* \right) \\ &= \sum_{(m,w) \in P} \left( 2 \sum_j x_{m,j}^* + 2 \sum_i x_{i,w}^* - \sum_{j \notin S} x_{m,j}^* - \sum_{i \notin S} x_{i,w}^* \right) \\ &\leq 4|P| - p_m - p_w, \end{aligned} \quad (8)$$

and similarly we also have

$$x^*(Q) \leq 4|Q| - q_m - q_w \quad (9)$$

and

$$x^*(R) \leq 4|R| - r_m - r_w. \quad (10)$$

To prove Lemmas 7 and 8, we use the following inequality, which is immediate from Constraints (1), (2), and (3) of  $LP(I)$ :

$$\begin{aligned} &x_{m,w}^* + \sum_{w \succ_m j} x_{m,j}^* + \sum_{m \succ_w i} x_{i,w}^* \\ &= \sum_j x_{m,j}^* + \sum_i x_{i,w}^* - \left( \sum_{j \succeq_m w} x_{m,j}^* + \sum_{i \succeq_w m} x_{i,w}^* - x_{m,w}^* \right) \\ &\leq 1. \end{aligned} \quad (11)$$

**Lemma 7**

$$x^*(P) \leq 3|P| - \sum_{(m,w) \in P} \sum_{w \succeq_m j, j \notin S} x_{m,j}^*.$$

*Proof* By the definition of  $x^*(P)$  and Constraints (1) and (2) of  $LP(I)$ ,

$$\begin{aligned} x^*(P) &\leq 2|P| + \sum_{(m,w) \in P} \left( \sum_{j \in S} x_{m,j}^* + \sum_{i \in S} x_{i,w}^* \right) \\ &= 2|P| + \sum_{(m,w) \in P} \left( \sum_{w \succ_m j, j \in S} x_{m,j}^* + \sum_{m \succ_w i, i \in S} x_{i,w}^* \right) \\ &\quad \text{(by Equations (6) and (7))} \\ &\leq 2|P| + \sum_{(m,w) \in P} \left( \sum_{w \succ_m j} x_{m,j}^* + \sum_{m \succ_w i} x_{i,w}^* \right) - \sum_{(m,w) \in P} \sum_{w \succ_m j, j \notin S} x_{m,j}^* \\ &\leq 3|P| - \sum_{(m,w) \in P} x_{m,w}^* - \sum_{(m,w) \in P} \sum_{w \succ_m j, j \notin S} x_{m,j}^* \quad \text{(by Inequality (11))} \\ &= 3|P| - \sum_{(m,w) \in P} \sum_{w \succeq_m j, j \notin S} x_{m,j}^*. \end{aligned}$$

□

**Lemma 8**  $x^*(R) \leq 3|R|$ .

*Proof* First, let us define  $s_{m,w}$  for  $(m, w) \in R$  as:

$$s_{m,w} = \sum_{m=w, i \in S} x_{i,w}^*.$$

Next, partition  $R$  into  $R_+$  and  $R_0$  such that  $R_+ = \{(m, w) \in R \mid s_{m,w} > 0\}$  and  $R_0 = \{(m, w) \in R \mid s_{m,w} = 0\}$ .

Now, consider  $(m, w) \in R_+$ . Since  $s_{m,w} > 0$ , there exists a man  $i \in S$  such that  $i =_w m$ . Since  $i$  is single in  $M$ ,  $f(i) = 3$  when  $i$  made the last proposal to  $w$ , and hence  $f(m) \geq f(i) = 3$  at the termination of the algorithm. Hence, during the execution of GSA-LP,  $m$  must have proposed to all the women in his list. Therefore, all of the women in his list are matched in  $M$ , and thus we have the following inequality (a):

$$\begin{aligned} x^*(R_+) &\leq 2|R_+| + \sum_{(m,w) \in R_+} \left( \sum_{j \in S} x_{m,j}^* + \sum_{i \in S} x_{i,w}^* \right) \\ &= 2|R_+| + \sum_{(m,w) \in R_+} \sum_{i \in S} x_{i,w}^* \\ &\leq 3|R_+|. \end{aligned}$$

Next, let us consider  $(m, w) \in R_0$ . By noting that men do not have ties, we have the following inequality (b):

$$\begin{aligned} x^*(R_0) &\leq 2|R_0| + \sum_{(m,w) \in R_0} \left( \sum_{j \in S} x_{m,j}^* + \sum_{i \in S} x_{i,w}^* \right) \\ &= 2|R_0| + \sum_{(m,w) \in R_0} \left( \sum_{w \succ_m j, j \in S} x_{m,j}^* + \sum_{m \succ_w i, i \in S} x_{i,w}^* \right) \\ &\quad \text{(by Equation (6) and } s_{m,w} = 0\text{)} \\ &\leq 2|R_0| + \sum_{(m,w) \in R_0} \left( \sum_{w \succ_m j} x_{m,j}^* + \sum_{m \succ_w i} x_{i,w}^* \right) \\ &\leq 3|R_0|. \quad \text{(by Inequality (11))} \end{aligned}$$

Finally, by inequalities (a) and (b), we have  $x^*(R) = x^*(R_+) + x^*(R_0) \leq 3|R_+| + 3|R_0| = 3|R|$ . □

#### 4.5 Proof of Lemma 3

Here, we define some more quantities to simplify later expressions:

$$\begin{aligned}\pi &= \sum_{(m,w) \in P} \sum_{w \succeq_m j, j \notin S} x_{m,j}^*, \quad \bar{\pi} = \sum_{(m,w) \in P} \sum_{j \succ_m w, j \notin S} x_{m,j}^*, \\ \bar{\alpha} &= \sum_{(m,w) \in P} \sum_{j \succ_m M_{opt}(m), j \notin S} x_{m,j}^*, \quad \text{and} \quad \bar{\beta} = \sum_{(m,w) \in Q} \sum_{j \succ_m M_{opt}(m), j \notin S} x_{m,j}^*\end{aligned}$$

Note that  $p_m = \pi + \bar{\pi} = \alpha + \bar{\alpha}$  and  $q_m = \beta + \bar{\beta}$ . (Recall that  $\alpha$  and  $\beta$  are defined in Sec. 4.2.)

Then, we show the following useful lemma.

**Lemma 9**  $|P| - \bar{\pi} \leq \alpha + \beta + r_m$ .

*Proof* It is easy to see that

$$\begin{aligned}|P| - \bar{\pi} &= \sum_{(m,w) \in P} \left( 1 - \sum_{j \succ_m w, j \notin S} x_{m,j}^* \right) \\ &= \sum_{(m,w) \in P} \left( 1 - \sum_{j \succ_m w} x_{m,j}^* \right) \\ &\leq \sum_{(m,w) \in P} \sum_{i \succeq_w m} x_{i,w}^*,\end{aligned}$$

where the second equality is from Equation (6) and the inequality is from Constraint (3) of  $LP(I)$ . Therefore we want to show that

$$\begin{aligned}\sum_{(m,w) \in P} \sum_{i \succeq_w m} x_{i,w}^* &\leq \sum_{(m,w) \in P} \sum_{M_{opt}(m) \succeq_m j, j \notin S} x_{m,j}^* \\ &\quad + \sum_{(m,w) \in Q} \sum_{M_{opt}(m) \succeq_m j, j \notin S} x_{m,j}^* + \sum_{(m,w) \in R} \sum_{j \notin S} x_{m,j}^*.\end{aligned}$$

The lefthand side of the inequality is the sum over all  $x_{i,w}^*$  such that (a)  $(m,w) \in P$ , (b)  $i \succeq_w m$ , and (c)  $x_{i,w}^* > 0$ . To prove the lemma, we show that every such  $x_{i,w}^*$  appears in the righthand side.

Let us consider a triple  $(m, i, w)$  that satisfies (a), (b), and (c). If  $i \in S$ , then  $x_{i,w}^* = 0$  by (a), (b), and Equations (6) and (7), which contradicts (c). Hence,  $i \notin S$  and so  $(i, M(i)) \in P \cup Q \cup R$ .

From (a),  $m \succ_w M_{opt}(w)$  by Lemma 5(ii). By using this and (b), we have (d)  $i \succeq_w m \succ_w M_{opt}(w)$ . Since  $w$  includes  $i$  in the list,  $i$  includes  $w$  in the list, and hence,  $M_{opt}(i)$  must exist and  $M_{opt}(i) \succeq_i w$ , as otherwise,  $(i, w)$  blocks  $M_{opt}$  by (d). It can be verified that such  $x_{i,w}^*$  appears in the righthand side, that is, if  $(i, M(i))$  is in  $P$ ,  $Q$ , and  $R$ , respectively, then  $x_{i,w}^*$  appears in the first, the second, and the third terms, respectively, of the righthand side.  $\square$

It is not hard to see that  $p_m + q_m + r_m = p_w + q_w + r_w$  and  $\bar{\pi} \leq p_m$  by definition and  $\bar{\pi} \leq \bar{\beta} \leq q_m$  by Lemma 6. Then, the following inequality is immediate by Inequalities (8), (9), and (10).

$$\begin{aligned}
 & x^*(P) + x^*(Q) + x^*(R) \\
 & \leq 4|P| + 4|Q| + 4|R| - 2(p_m + q_m + r_m) \\
 & \leq 4|P| + 4|Q| + 4|R| - 4\bar{\pi} - 2r_m \quad (\text{because } \bar{\pi} \leq p_m \text{ and } \bar{\pi} \leq q_m) \\
 & \leq 4|Q| + 4|R| + 4(\alpha + \beta + r_m) - 2r_m \quad (\text{by Lemma 9}) \\
 & = 2|P| + 2|Q| + 4|R| + 4\alpha + 4\beta + 2r_m \quad (\text{because } |P| = |Q|) \\
 & \leq 2|P| + 2|Q| + 6|R| + 4\alpha + 4\beta. \quad (\text{because } r_m \leq |R|)
 \end{aligned}$$

□

#### 4.6 Proof of Lemma 4

First, we show the following useful lemma.

##### Lemma 10

$$\sum_{(m,w) \in P} \sum_{j \succ_m M_{opt}(m)} x_{m,j}^* + \sum_{(m,w) \in Q} \sum_{i \succeq_w m} x_{i,w}^* \geq |Q|.$$

*Proof* Consider an augmenting five-path of Fig. 1(b). We have

$$\sum_{j \succ_{m_p} w_q} x_{m_p,j}^* + \sum_{i \succeq_{w_q} m_q} x_{i,w_q}^* = \left( \sum_{j \succeq_{m_p} w_q} x_{m_p,j}^* - x_{m_p,w_q}^* \right) + \sum_{i \succeq_{w_q} m_p} x_{i,w_q}^* \geq 1,$$

where the equality is from  $m_p =_{w_q} m_q$  (Lemma 5(iv)) and the inequality is from Constraint (3) of  $LP(I)$  (note that  $(m_p, w_q) \in A$ ). By summing up this inequality for all of the augmenting five-paths, we obtain this lemma. □

By Lemma 5(iii), we have  $M_{opt}(m) \succ_m w$  for  $(m, w) \in P$ . Therefore, by using Equations (6) and (7), Lemma 10 can be written as:

**Corollary 1**  $\bar{\alpha} + q_w \geq |Q|$ .

Then the following inequality follows from Lemmas 7 and 8, and Inequality (9).

$$\begin{aligned}
 & x^*(P) + x^*(Q) + x^*(R) \\
 & \leq 3|P| + 4|Q| + 3|R| - \pi - q_m - q_w \\
 & = 3|P| + 4|Q| + 3|R| - \pi - \bar{\beta} - \beta - q_w \quad (\text{because } q_m = \beta + \bar{\beta}) \\
 & \leq 3|P| + 4|Q| + 3|R| - \pi - \bar{\pi} - \beta - q_w \quad (\text{because } \bar{\pi} \leq \bar{\beta} \text{ by Lemma 6}) \\
 & = 3|P| + 4|Q| + 3|R| - \alpha - \bar{\alpha} - \beta - q_w \quad (\text{because } \pi + \bar{\pi} = \alpha + \bar{\alpha}) \\
 & \leq 3|P| + 3|Q| + 3|R| - \alpha - \beta. \quad (\text{by Corollary 1})
 \end{aligned}$$

□



## 5 Integrality Gaps

In this section, we consider the limitation of our method. In Sec. 5.1, we show that the integrality gap between our IP formulation and LP relaxation is at least  $1 + \frac{1}{e}$  ( $> 1.3678$ ). This does not immediately imply a lower bound of GSA-LP, but may be considered as some barometer to show its limit. Also, in Sec. 5.2, we use a similar approach to show that the integrality gap is at least  $1.5 - o(1)$  if we use the same IP formulation for the general MAX SMTI (i.e., ties on both sides). Therefore, generalizing Király's GSA2 [15] (Király's 5/3-approximation algorithm for the general MAX SMTI) or McDermid's algorithm [18] using an optimal LP solution seems difficult unless a fundamentally new idea is introduced.

### 5.1 One-Sided Ties

Consider the following instance  $I_2$  of MAX SMTI.

$$\begin{array}{ll} m_1: & w_1 \cdots w_k w_{k+1} & w_1: & (m_1 \cdots m_k) m_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ m_k: & w_1 \cdots w_k w_{2k} & w_k: & (m_1 \cdots m_k) m_{2k} \\ m_{k+1}: & w_1 & w_{k+1}: & m_1 \\ \vdots & \vdots & \vdots & \vdots \\ m_{2k}: & w_k & w_{2k}: & m_k \end{array}$$

An optimal solution  $M_{opt} = \{(m_i, w_i) \mid i = 1, \dots, k\}$  has size  $k$ , and hence the optimal value of  $IP(I_2)$  is  $k$ . However  $LP(I_2)$  has the following solution  $x$ .

- $x_{m_i, w_j} = \frac{1}{k} \left(1 - \frac{1}{k}\right)^{j-1}$  for  $(i, j) \in \{1, \dots, k\} \times \{1, \dots, k\}$ ,
- $x_{m_i, w_{k+i}} = \left(1 - \frac{1}{k}\right)^k$  for  $i = 1, \dots, k$ ,
- $x_{m_{k+i}, w_i} = 1 - \left(1 - \frac{1}{k}\right)^{i-1}$  for  $i = 1, \dots, k$ .

The value of the objective function for  $x$  is

$$\begin{aligned} \sum_i \sum_j x_{m_i, w_j} &= k \sum_{j=1}^k \frac{1}{k} \left(1 - \frac{1}{k}\right)^{j-1} + k \left(1 - \frac{1}{k}\right)^k + \sum_{i=1}^k \left(1 - \left(1 - \frac{1}{k}\right)^{i-1}\right) \\ &= k + k \left(1 - \frac{1}{k}\right)^k. \end{aligned}$$

Hence the integrality gap is

$$\sum_i \sum_j x_{m_i, w_j} / |M_{opt}| = 1 + \left(1 - \frac{1}{k}\right)^k.$$

This value approaches  $1 + \frac{1}{e}$  ( $> 1.3678$ ) as  $k$  goes to infinity.

## 5.2 Two-Sided Ties

We give an instance  $I_3$  with two-sided ties such that the integrality gap is  $1.5 - o(1)$  for  $IP(I_3)$  and  $LP(I_3)$ .

$$\begin{array}{ll}
 m_1: & (w_1 \cdots w_k) w_{k+1} & w_1: & (m_1 \cdots m_k) m_{k+1} \\
 \vdots & \vdots & \vdots & \vdots \\
 m_k: & (w_1 \cdots w_k) w_{2k} & w_k: & (m_1 \cdots m_k) m_{2k} \\
 m_{k+1}: & w_1 & w_{k+1}: & m_1 \\
 \vdots & \vdots & \vdots & \vdots \\
 m_{2k}: & w_k & w_{2k}: & m_k
 \end{array}$$

An optimal solution  $M_{opt} = \{(m_i, w_i) \mid i = 1, \dots, k\}$  has size  $k$  and hence the optimal value of  $IP(I_3)$  is  $k$ . However  $LP(I_3)$  has this solution  $x$ :

- $x_{m_i, w_j} = 1/(2k-1)$  for  $(i, j) \in \{1, \dots, k\} \times \{1, \dots, k\}$ ,
- $x_{m_i, w_{k+i}} = 1 - k/(2k-1)$  for  $i = 1, \dots, k$ ,
- $x_{m_{k+i}, w_i} = 1 - k/(2k-1)$  for  $i = 1, \dots, k$ .

The value of the objective function for  $x$  is

$$\sum_i \sum_j x_{m_i, w_j} = k^2 \cdot \frac{1}{2k-1} + 2k \left(1 - \frac{k}{2k-1}\right) = \frac{3k-2}{2k-1} k.$$

Hence the integrality gap is

$$\sum_i \sum_j x_{m_i, w_j} / |M_{opt}| = \frac{3k-2}{2k-1} k \cdot \frac{1}{k} = \frac{3}{2} - \frac{1}{2(2k-1)} = \frac{3}{2} - o(1).$$

## 6 Concluding Remarks

In this paper, we presented a new upper bound of 25/17 on the approximation ratio of MAX SMTI with one-sided ties, by extending Király's and McDermid's algorithms using IP and LP relaxation. An apparent future work is to further improve this upper bound. One possibility is to improve Lemmas 3 and 4 because we believe that they are not tight. Another apparent future work is to extend this method to the general MAX SMTI, which may require some new ideas as mentioned in Sec. 5.

**Acknowledgements** The authors would like to thank the anonymous reviewers for their valuable comments.

## References

1. A. Abdulkadiroğlu, P. A. Pathak, A. E. Roth, and T. Sönmez, “The Boston public school match,” *American Economic Review P&P*, Vol. 95, No. 2, pp. 368–371, 2005.
2. A. Abdulkadiroğlu, P. A. Pathak, A. E. Roth, and T. Sönmez, “Changing the Boston school choice mechanism: strategy-proofness as equal access,” manuscript, 2006.
3. D. Gale and L. S. Shapley, “College admissions and the stability of marriage,” *Amer. Math. Monthly*, Vol. 69, pp. 9–15, 1962.
4. D. Gale and M. Sotomayor, “Some remarks on the stable matching problem,” *Discrete Applied Mathematics*, Vol. 11, pp. 223–232, 1985.
5. D. Gusfield and R. W. Irving, *The Stable Marriage Problem: Structure and Algorithms*, MIT Press, Boston, MA, 1989.
6. M. M. Halldórsson, K. Iwama, S. Miyazaki, and H. Yanagisawa, “Randomized approximation of the stable marriage problem,” *Theoretical Computer Science*, Vol. 325(3), pp. 439–465, 2004.
7. M. M. Halldórsson, K. Iwama, S. Miyazaki, and H. Yanagisawa, “Improved approximation results for the stable marriage problem,” *ACM Transactions on Algorithms*, Vol. 3(3), Article No. 30, 2007.
8. R. W. Irving, “Stable marriage and indifference,” *Discrete Applied Mathematics*, Vol. 48, pp. 261–272, 1994.
9. R. W. Irving and D. F. Manlove, “Approximation algorithms for hard variants of the stable marriage and hospitals/residents problems,” *Journal of Combinatorial Optimization*, Vol. 16(3), pp. 279–292, 2008.
10. R. W. Irving, D. F. Manlove, and G. O’Malley, “Stable marriage with ties and bounded length preference lists,” *Journal of Discrete Algorithms*, Vol. 7(2), pp. 213–219, 2009.
11. K. Iwama, D. F. Manlove, S. Miyazaki, and Y. Morita, “Stable marriage with incomplete lists and ties,” In *Proc. ICALP*, LNCS 1644, pp. 443–452, 1999.
12. K. Iwama, S. Miyazaki and K. Okamoto, “A  $(2 - c \log N/N)$ -approximation algorithm for the stable marriage problem,” *IEICE Transactions*, Vol. 89-D(8), pp. 2380–2387, 2006.
13. K. Iwama, S. Miyazaki, N. Yamauchi, “A  $(2 - c \frac{1}{\sqrt{N}})$ -Approximation Algorithm for the Stable Marriage Problem,” *Algorithmica*, Vol. 51(3), pp. 342–356, 2008.
14. K. Iwama, S. Miyazaki, and N. Yamauchi, “A 1.875-approximation algorithm for the stable marriage problem,” In *Proc. SODA*, pp. 288–297, 2007.
15. Z. Király, “Better and simpler approximation algorithms for the stable marriage problem,” *Algorithmica*, Vol. 60, No. 1, pp. 3–20, 2011.
16. S. Khot and O. Regev, “Vertex cover might be hard to approximate to within  $2 - \epsilon$ ,” *Journal of Computer and System Sciences*, Vol. 74(3), pp. 335–349, 2008.
17. D. F. Manlove, R. W. Irving, K. Iwama, S. Miyazaki, and Y. Morita, “Hard variants of stable marriage,” *Theoretical Computer Science*, Vol. 276(1–2), pp. 261–279, 2002.
18. E. J. McDerimid, “A  $3/2$ -approximation algorithm for general stable marriage,” In *Proc. ICALP*, LNCS 5555, pp. 689–700, 2009.
19. A. E. Roth and E. Peranson, “The redesign of the matching market for American physicians: some engineering aspects of economic design,” *American Economic Review*, Vol. 89, No. 4, pp. 748–780, 1999.
20. A. E. Roth, U. G. Rothblum, and J. H. V. Vate, “Stable matchings, optimal assignments, and linear programming,” *Mathematics of Operations Research*, Vol. 18(4), pp. 803–828, 1993.
21. C.-P. Teo and J. Sethuraman, “The geometry of fractional stable matchings and its applications,” *Mathematics of Operations Research*, Vol. 23(4), pp. 874–891, 1998.
22. C.-P. Teo, J. Sethuraman, and W. P. Tan, “Gale-Shapley stable marriage problem revisited: strategic issues and applications,” In *Proc. IPCO*, LNCS 1610, pp. 429–438, 1999.
23. H. Yanagisawa, “Approximation Algorithms for Stable Marriage Problems,” Ph. D. Thesis, Kyoto University, 2007.